

PCT

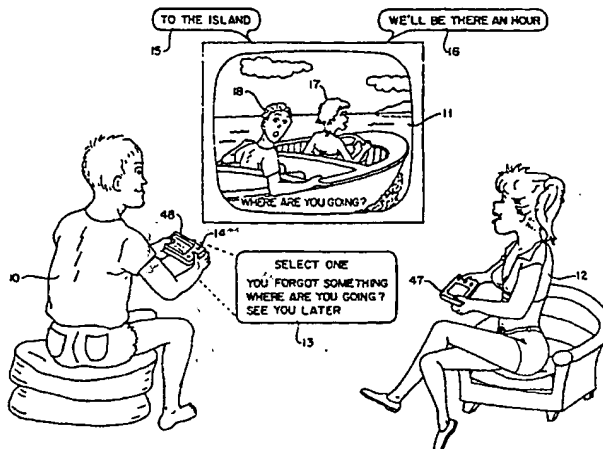
WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : A63B 9/24	A1	(11) International Publication Number: WO 94/27677 (43) International Publication Date: 8 December 1994 (08.12.94)
(21) International Application Number: PCT/US93/05237 (22) International Filing Date: 26 May 1993 (26.05.93) (30) Priority Data: 08/066,204 25 May 1993 (25.05.93) US (71)(72) Applicant and Inventor: BEST, Robert, Macandrew [US/US]; 16016 Ninth Avenue N.E., Seattle, WA 98155 (US). (74) Agents: GRAYBEAL, John, O. et al.; Graybeal Jackson Haley & Johnson, 777-108th Avenue N.E. #2460, Bellevue, WA 98004 (US).	(81) Designated States: AT, AU, BB, BG, BR, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SK, UA, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published With international search report. With amended claims.	

(54) Title: TALKING VIDEO GAMES



(57) Abstract

This invention consists of methods of simulating voice conversations with talking animated characters (17, 18) on a television or video screen (11). The animated characters (17, 18) talk to each other and to humans (10, 12). Each human holds a light-weight controller (47, 48) that has push buttons (14) next to a display of variable phrases or sentences (13) for each human's side of the dialog. This dialog includes alternative words (13, 26) for a human player to say to a character or for a character to say (15, 16, 20, 23, 27) or actions (Figs. 3, 7, 10) for a character to do. A human (10, 12) responds to what a character (17, 18) says or does by pressing a button (14) next to a selected phrase. An animated character then vocally responds (15 in Fig. 1) to the selected phrase as if it had been spoken by a human or as if it were the character's own words. Each scene (64) branches to subsequent scenes (65, 66, 67) and within each scene there are several branching dialog sequences (60, 61, 62, 69, 63, 68).

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

-1-

TALKING VIDEO GAMES

Technical Field

This invention relates to video games, animated cartoons, and picture/sound synchronization.

5 Background Art

We are all born with a desire to talk and to be talked to. Listening to other people talk and thereby sharing their emotional experiences is also a desire we are born with and this desire has been partly satisfied
10 by motion picture film and television in which voice sounds are now essential. Until recently, voice sounds were seldom used in video games or were used in an optional manner. People have a desire to participate in voice conversations with other people, but this
15 desire is not satisfied by current video games.

Adding voice sounds to conventional video games is not enough to simulate a face to face voice conversation. A talking video game called Thayer's Quest was attempted in 1984 and was played from an
20 analog laser-readable video disc. One of the reasons for the commercial failure of Thayer's Quest was that each spoken sentence was programmed to accompany only one sequence of video frames. Since the video was not compressed, the maximum amount of play time was limited
25 to about half an hour which was further reduced to a fraction of that by the branching story. Hence, only a few minutes of voice sounds were actually heard during the game. Whenever a human player saw a certain video character, the character usually spoke the same words.
30 This greatly reduced the entertainment value of the game. Another consequence of programming the audio and video to be inseparable, was that branching scenes were not distinguished from branching dialog. Talking video games will be much more entertaining if each scene has
35 a rich variety of possible dialog sequences.

-2-

Talking video games are disclosed in my U.S. Patents Nos. 4,305,131; 4,333,152; 4,445,187 and 4,569,026. These patents show only one talking animated character and one human player, although multiple characters and players were suggested in the texts. These patents also show push-button controllers that display alternative words to say to animated characters.

It is common practice in the video game art for stories to branch. It is also common practice for digital data of animated characters to be stored separately from background scenery and to generate each frame of an animated picture from both the background data and the character data to minimize the number of stored images.

It is well known for human players to input choices using any of a variety of input devices such as push buttons, rotatable knobs, pressure sensitive membrane, proximity sensitive pads or screen overlay, light pen, light sensitive gun, joy stick, mouse, track ball, moving a cursor or crosshairs or scrolling through highlighted options, speech recognition, etc.

In the prior art, each choice by the human can be immediately followed by a synthesized voice or digitized voice recording that speaks the words selected by the human player, so the human will quickly adjust to the fact that the spoken words he hears for his side of the dialog are initiated by his fingers rather than his vocal cords.

The prior art also includes methods for generative video images of moving lips and facial expressions on a talking head or other animated character. See for example, U.S. Patent No. 4,884,972 issued December 5, 1989 to Elon Gasper who contemplates use in video games.

Drawing Figures

FIG. 1 illustrates an animated cartoon talking game with two human players who take turns selecting words to say to one or two animated characters on a TV screen who then respond to the words.

-3-

FIG. 2 illustrates an animated talking game with one human player and two animated cartoon characters on a TV screen who talk to the human and to each other.

5 FIG. 3 is an animated cartoon illustrating a kissing scene in which one character's unspoken thoughts are sounded.

FIG. 4 is an animated cartoon illustrating a ledge-hanging scene in which a talking cartoon character requires immediate action by the human player.

10 FIG. 5 is an animated sequence illustrating branching dialog in which two cartoon characters respond to sentences selected by a human player by talking to the human player and to each other.

15 FIG. 6 is an animated sequence illustrating branching dialog in which one animated character responds to a sentence selected by a human player and refers to a second character (not shown).

20 FIG. 7 is an animated sequence showing hands simulating the hands of a human player performing an action.

FIG. 8 is a flow diagram illustrating scene branching and dialog branching within each scene.

25 FIG. 9 is a pictorial view of a video game system having a CD-ROM disc drive connected to a television or video monitor with auxiliary devices for displaying prompting words.

30 FIG. 10 is an animated sequence illustrating both dialog branching and scene branching in which one character is shown helping the other two characters get out of a dangerous situation.

FIG. 11 is an apparatus block diagram illustrating flow of data and method steps to generate a cyclical game play with two animated characters.

-4-

Description of Preferred Embodiments

The characters in role-playing video games are of two kinds: player-controlled characters and non-player characters. A player-controlled character is a human player's animated counterpart and does what the human player chooses to have him do. Non-player characters are not directly controlled by a human player, but can be indirectly influenced by a human player, either by responding to an action selected by the human player or by responding to what a player-controlled character does or says. Non-player characters should be programmed to say many different alternative things depending on what a player or player-controlled character has just "said". A scene may repeat several times with a non-player character saying something different each time and the alternative responses for each human player being different each time a scene is repeated.

Hand-held controllers are preferable to displaying menus on the video screen, because in two-person games the menu displayed to one player may be different from the other player's menu. Part of the fun comes from a player not knowing what choices the other player has. Also, video screen menus disrupt the illusion that a player is talking with a real video character. Hence, the preferred embodiment of my invention makes use of hand-held menu display controllers with a push-button or equivalent for each menu item. The invention also makes use of echo voices that repeat the selected words so that other human players will know what is being said.

Referring to FIG. 1, in a typical embodiment of this invention a video game system generates for display on a video screen 11 an animated picture sequence to two human game players. Human players 10 and 12 take turns selecting a phrase or sentence to "say" to a character or characters on a video screen who then talk back responsively. In this example it is player 10's turn. Player 12's hand-held

-5-

controller is blank, indicating that she cannot say anything to a character at this moment. The video screen shows two animated characters traveling in a boat or other vehicle. The characters are non-player characters that talk directly to human players in this example. Each human player holds a hand-held controller with three push buttons next to a liquid-crystal display 13, shown enlarged in FIG. 1 for clarity.

As the boat scene begins, the video game system displays two or three or more alternative responses on display 13 or other display device. While player 10 is deciding which button to press, a linking picture sequence (not shown) continues to show the same boat scene with one character looking back from the boat as if looking at player 10 standing on the dock. Player 10 selects one of the displayed responses (in this example "WHERE ARE YOU GOING?") which is then echoed as voice sounds or is displayed as sub-titles on the video screen so that player 12 will know what player 10 has just said to an animated character. This character 18 answers responsively with voice sounds: "TO THE ISLAND." after which a second character 17 responds with "WE'LL BE THERE AN HOUR." in response to what character 18 just said.

In the preferred embodiment, the game system generates a voice to echo the words selected by player 10 so that player 12 hears what player 10 has "said" to the animated character. When it is player 12's turn, player 12's hand-held controller will then display alternative words (not shown in FIG. 1) that she can "say" in response to what animated character 17 or 18 has said. Thus a three-way or four-way dialog can be simulated.

Before each game begins, display 13 may show identifying information for each player-controlled character so that each human player may select which character he or she wants to play. Alternatively, each player-controlled character may be shown on video screen 11 and the question asked "Who wants to play this

-6-

character?" for each character shown. The game system then records which hand-held controller responds to the question, so later that character's words will be shown on display 13 only for the player who is playing that character (in this example player 10). A human player may also choose to play a player-controlled character that need not always be shown on the video screen, but who may be off-screen (for example, standing on the dock in FIG. 1), and can carry on a dialog with on-screen characters.

Referring to FIG. 2, in another embodiment of the invention, a video game system generates for display on a video screen 11 an animated picture sequence having two or more animated talking characters 17 and 18. In this example, character 17 is a player-controlled character that human player 12 controls. Player 12 plays the role of character 17 and can talk to character 18 and other characters in the game through character 17. Player 12 holds a hand-held controller 47 with three push buttons next to a liquid-crystal display 13, shown enlarged in FIG. 2 for clarity. The game system displays three alternative responses on display 13. Player 12 selects one of the displayed responses (in this example "KISS ME AGAIN") with push button 14. The echo voice sounds 15 for character 17 then repeat the words selected from display 13. The words are directed at character 18 whose voice sounds 16 then respond to character 17.

The two animated characters 17 and 18 may respond in reverse sequence, that is, the non-player character 18 may say his line first so that the player-controlled character 17 can respond as ordered by human player 12. For example, after player 12 selects "KISS ME AGAIN", character 18 may say "WHAT DO YOU WANT TO DO?" a prerequisite sentence that is not one of the displayed alternatives, then character 17 can respond with "KISS ME AGAIN" which is responsive to what character 18 has just said.

Echo voices or sub-titles may also be used to express unspoken thoughts or the thoughts of non-speaking beings

-7-

such as babies or animals or inanimate objects. Cloud balloon 19 in FIG. 2 represents an unspoken thought of character 18 which is sounded (with mouth not moving) in response to spoken sentence 15 of character 17. Voice sounds for unspoken thoughts may be electronically altered to indicate to players that a voice is not a normal spoken voice. For example, unspoken thoughts can be given a hollow or tremulous sound or a whispering sound by electronically or digitally editing voice sound data before converting to audio.

Referring to FIG. 3, when characters 17 and 18 perform a competitive or cooperative activity such as kissing, one of the characters may speak (with moving mouth) or think unspoken thoughts (sounded with unmoving mouth) as in cloud balloon 29, responsive to the action being shown or to what was said or done in the prior sequence shown in FIG. 2 or in response to a phrase selected from display 13.

Referring to FIG. 4, in another embodiment of this invention, a video game system generates an animated picture sequence showing a character 30 hanging by his fingers from a ledge on the outside of a building. His friends on the roof have thrown him a rope which is not long enough to reach character 30. This ledge-hanging scene is followed by the next scene shown in FIG. 5.

Referring to FIG. 5, picture sequences 21, 25, and 28 are parts of one roof scene in which two talking characters discuss how to rescue the character in FIG. 4. One or two human players participate in the conversation by "saying" words or phrases or sentences to the animated characters who then answer responsively and ask questions or make remarks that lead to the next input by a player. The alternatives shown on display 22 are suggested solutions to the problem posed in sentence 20. When the human player 10 presses button 14 next to "CALL EMERGENCY", character 17 responds by asking

-8-

question 23 to character 18 who responds with question 24 directed at the human player. Question 24 is also accompanied by display of a new menu of alternative actions 26. When a player presses the button next to
5 "SLIDE DOWN THE ROPE", character 17 comments on this choice of action with sentence 27 in sequence 28. Thus a simulated verbal dialog can continue through several exchanges of words within the same scene.

Referring to FIG. 6, in another embodiment of this
10 invention a video game system generates animated picture sequence 31 and 32. The picture sequence has one scene showing a burning car in the background and a frightened woman 17 in the foreground. During part 31 of the sequence the woman's voice 15 says "PLEASE HELP ME! MY
15 MOTHER IS TRAPPED IN THE CAR!" The game then displays two or more alternative actions on hand-held display 13 or on the video screen or other display device. The human player 10 selects one of the displayed responses (such as "PRY THE DOOR OPEN") and presses the corresponding push
20 button 14 or equivalent. While the player is deciding which button to press, a linking picture sequence (not shown) continues to show the same scene with the woman looking anxiously at the player. When the player selects a response, part 32 of the animated sequence continues
25 showing the same burning car scene with the woman's voice 16 answering responsively: "I TRIED TO, BUT I CAN'T GET IT OPEN".

Selecting a response by pushing a button 14 can result in a change to a different scene, but in this FIG. 6
30 example the scene remains the same while the branching dialog continues. Each of the three alternative responses on display 13 will result in a different answer by animated character 17. The animated sequences 31 and 32 are generated by the video game system by overlaying three
35 moving pictures: (1) the background sequence showing the burning car with flames continually flickering, (2) a sprite or mosaic of sprites showing the woman character 17

-9-

moving in a natural manner against the background and (3) mouth and other facial sprites selected (by table lookup) by the game system approximately lip-synchronized with voice sounds 15 and 16.

5 Since the player's response time is variable, sequence 31 or 32 with a closed mouth sprite should cycle continually until a button 14 is pressed or until a time limit is reached, at which time a prompting picture sequence with words such as "Hurry! Hurry!" may be sounded
10 and displayed with moving mouth sprites. The burning car can be any object such as a damaged bus, airplane, boat, or building that provides a danger situation.

15 Referring to FIG. 7, the game system generates an animated sequence showing one or more hands 36 simulating the hands of a human player performing an action. In this example, animated sequence 33 shows a hand holding pipe 35 or a fire extinguisher or other such tool while the human player whose real hand 10 is illustrated pressing button 14 selects one of the
20 alternative actions shown on display 13. Note that display 13 in FIG. 7 shows alternative actions which would result in scene changes, but display 13 in FIG. 6 shows alternative words or phrases to say to animated character 17 in a common scene followed by a scene
25 change. In FIG. 7, pressing button 14 to select "PRY THE DOOR OPEN" results in picture sequence 34 showing hands 36 prying open the car door with pipe 35.

30 Referring to FIG. 8, a flowchart illustrates the distinction between branching dialog and branching scenes. For example, at dialog branch point 60 a player is presented with two or more alternatives to choose from which correspond to the menu of verbal expressions displayed on hand-held controller 48 in Figures 1-7. Each of the alternative selections will
35 result in a different verbal response 61 or 62 by one or more of the animated characters, but does not result in a scene change in this instance. However, the

-10-

alternatives at branch point 63 in FIG. 8 will result in a scene change, either to scene 65 or to scene 66. Branch point 69 may result in either a scene change to scene 67 or a dialog branch depending on a player's selection. Dialog branching in each scene provides a large variety of possible dialog sequences without duplicating the animated picture data for each dialog branch. New animated picture data is needed only whenever branching to a new scene occurs. If player 10 in FIG. 6 selects the first of the three alternative actions shown on display 13, the scene changes to the pipe scene in FIG. 7 which corresponds to response 68 in FIG. 8. Response 68 includes voice sounds 16 of character 17 in FIG. 6 followed by a scene change and actions by a second character 36 in FIG. 7.

Referring to FIG. 9, video game system 42 is connected by cable 44 to a television or video monitor 11 and to one or more hand-held control units 48 and 47 or portable game system 46, each having three push buttons 14 next to a liquid-crystal display 13 showing a variable menu of verbal expressions. Infra-red or radio signals may be used instead of cables. System 42 includes a disk reader that reads digital data from a CD-ROM disk 43, or write-once disk or card or other medium containing digital data from which system 42 generates animated picture signals. Disk 43 also contains compressed audio for voice and other sounds, synchronizing data, and digitized words for display on units 48 and 47. Portable game system 46 with appropriate ROM program cartridge may substitute for display units 48 and 47. Cable 45 connects game system 42 to display unit 48, 46 or 47 and transmits alternative prompting words or other verbal expressions for display. Cable 45 also transmits push-button 14 selection signals to system 42.

If portable game system 46 has only one or two available push-buttons, selection from among responses shown on display 13 may be done as shown in FIG. 7 by

-11-

pressing push-button 14 multiple times to position a pointer 49, cursor, or other such indicator next to (or on) a selected sentence.

Referring to FIG. 10, in one embodiment of this invention a video game system generates on a video screen 11 several sequences of animated pictures 21, 25, and 28 showing three talking characters engaged in a rescue activity. The characters have distinctive voices. As sequence 21 begins, characters 18 and 30 are in a dangerous situation and character 17 is trying to rescue them. Character 17 speaks words represented in voice balloon 20 commenting on the rescue activity. A human player 12 in this example holds a hand-held controller 47 with three push buttons 14 positioned next to corresponding menu items on a liquid-crystal display 13. Two or more alternative response sentences or phrases are shown on display 13, each being associated with one of the three push buttons 14. The hand of human player 12 is shown pressing one of the buttons 14 that selects one of the alternative sentences to be spoken by character 30.

While human player 12 is deciding which button to press, a linking picture sequence (not shown) continues to show the same scene with the characters trying to pull up on the rope or trying to climb up the rope.

When player 12 selects one of the displayed alternative response sentences that calls for an action shown in a different scene, branching to the next scene occurs beginning in this example with animated picture sequence 25. The selected sentence is sounded in the voice of player-controlled character 30 who repeats the words selected by human player 12. These voice sounds are represented in balloon 23 (in this example "GRAB THE ROPE TOM"). Tom is character 18 who is shown in sequence 21 hanging dangerously onto an ankle of character 30. After character 30 says the selected sentence in balloon 23, character 18 grabs hold of the

-12-

rope and then responds in his voice with the words "I'VE GOT IT" which are the preprogrammed words for the selected sentence. These voice sounds of character 18 are represented in balloon 24. The actions of
5 character 18 are the preprogrammed actions for the selected sentence. If human player 12 had selected a different one of the alternative sentences, a different scene would result and character 18 would have behaved differently and spoken different response words.

10 The video game system next generates an animated picture sequence showing character 18 climbing the rope. This climbing sequence is not shown in the drawings. When character 18 reaches character 17 in sequence 28, character 17 says in her voice "HOLD MY
15 ARM" (in balloon 27) and the rescue activity proceeds. Thus, the animated characters perform or attempt to perform actions and talk about these actions while both the actions and voice sounds are controlled directly or indirectly by a human player or players. The words
20 spoken by character 17 may also be selected by a human player from a second menu (not shown) of alternative sentences for her to say. This human player may be a second player 10 as shown in FIG. 5.

25 After the end of sequence 28, the game system generates another animated sequence, then another sequence, then another sequence, etc. in which a human player controls, directly or indirectly, character actions and voice sounds as described above for FIG. 10.

30 Referring to FIG. 11 which is a block diagram of a special-purpose apparatus for performing some of the video game illustrated in FIG. 1, the apparatus performs steps 1 through 7 in sequence, so that the animated picture sequence on the television or video
35 screen 11 alternates between two characters in this example. In step 1, animated picture generator 51 generates a video or RF signal 44 for display on video

-13-

screen 11 of an animated picture from digitized animated character data stored in memory 70 of a first talking character. In step 2, hand-held controller 48 displays on the liquid-crystal display 13 several alternative sentences or other verbal expressions from data read via line 72 from memory 52 storing the alternative sentences. In step 3, one of the push buttons 14 generates a selection signal (indicated by one of lines 53) selecting a sentence from among the alternative sentences in memory 52 and also selects the corresponding (indicated by dotted lines) digitized voice recordings from memory 57 and 58. In step 4, voice signal generator 54 generates a signal 41 for output as first voice sounds from the selected digitized voice recording (via line 73) in memory 57 that expresses or responds to the words displayed on controller 48. In step 5, animated picture generator 55 generates a signal 44 for display on video screen 11 of an animated picture from digitized animated character data stored in memory 71 of a second talking character. In step 6, voice signal generator 56 generates a signal 41 for output as second voice sounds from the selected digitized voice recording (via line 74) in memory 58 that respond to the words generated by voice signal generator 54. In step 7, a new set of alternative sentences from disk or memory 59 is loaded into memory 52, 57 and 58 and the sequence of steps begins again with step 1.

To allow each background scene to be used with different animated characters who can move around against the background scene, the digital animation data for the background scene should be stored separately from the digital animation data for each character. Similarly, to allow each character to say many different sentences without a scene change, the digitized voice data should be independent of the animation data. In the preferred embodiment, animated character video, voice sound sequences and guidance word sequences are independently

-14-

generated from separately stored digital data. Dialog data that is not used in one scene may be used later in a different scene with the same or different characters. The voice data may consist of sequences of codes or compressed digital recordings of words, phrases, word segments or phonemes in several distinctive voices so that each character can speak thousands of preprogrammed words or sentences. Similarly, the digital data for each animated character's body may be stored separately from sprite data for moving lips, facial expressions, and gestures, so that each character and its distinctive voice can be lip-synchronized with different mouth movements depending on the dialog. The digital data for each animated character may also combine body images, lips, facial expressions, hand gestures and voice sounds.

When a human player presses one of the buttons 14 (FIG. 1, 2, 5, 6 or 7) the game system may generate a voice sound speaking the selected sentence as a substitute for the player's side of the dialog. The animated character then "responds" as if the generated voice sounds had been spoken by the human player. Because the player selects the words which are actually sounded, he will quickly adjust to the fact that the spoken words he hears for his side of the dialog are initiated by his fingers rather than his vocal cords. This echo voice repeating the selected words is important for games with multiple human players so that each player will hear what each of the other players has "said" to on-screen characters. Pushing one of the buttons 14 selects both a simulated verbal response to the words previously spoken by an animated character and also selects a new dialog sequence that corresponds to the simulated verbal response shown on display 13. The selected dialog sequence includes the face and voice of the animated character speaking words which are responsive to the player's selected verbal response.

-15-

Alternatively, sub-titles such as illustrated on video screen 11 in FIG. 1 may be used instead of echo voices and be displayed on a video screen or on a hand-held controller as a substitute for the player's side of the dialog.

Each character can be an animated cartoon, digitized live action, analog live action, a sprite or the like, and be player controlled or not. The term "verbal expression" means any word, words, phrase, sentence, question, expletive, curse, keyword, combination of keywords, symbol, icon, or any meaningful human voice sound such as "huh?" or "hmmm" or laughter or scream. The word kissing is used herein to mean simulated touching of one animated character's mouth to another animated character's mouth or other body part.

No video game yet satisfies the basic human desire for people to speak and respond to us as individuals. Although the telephone provides a partial satisfaction of our desire to talk with other people, it is necessarily limited to living people who are willing to talk with us. Historical and imaginary people cannot talk with us and famous living people do not want to. Hence, there is a strong but unfulfilled human desire waiting to be satisfied by new technology. Talking animated characters will change the nature of video games as dramatically as talking pictures changed silent film. Talking video games will let human players talk with images of famous people (living, historical or imaginary) and with animal-like characters, and participate in simulated adventures and dramas and conversations with interesting characters and groups of characters who will talk to each player responsively. Talking video games that stir human emotions like dramatic films will have lasting appeal, because they will satisfy a basic human desire, the desire to talk with other people.

-16-

Although I have described the preferred embodiments of my invention with a degree of particularity, it is understood that the present disclosure has been made only by way of example and that equivalent steps and components may be substituted and design details changed without departing from the spirit and scope of my invention.

-17-

CLAIMS

1. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of two or more talking characters (17,18) for display on a video screen (11) and generates a signal (41) representing the respective voices of the talking characters,
 - 5
 - 10
 - 15
 - 20
 - 25
- a) digitally storing animated picture data (70,71) for first (18) and second (17) talking characters;
- b) said video system (42) generates (51,55) a signal (44) from said animated picture data (70,71) for display as animated pictures on said video screen (11);
- c) said video system (42) digitally stores first voice sound data (57) for said first talking character (18) and generates (54) a signal (41) from this first voice sound data (57) for output as first voice sounds;

-18-

d) said video system (42) transmits (72) a first set of alternative guidance words or other preprogrammed verbal expressions (52) in digital form to a first manually operated control device (48) which displays (13) the first set of alternative verbal expressions (52);

e) one of said first set of verbal expressions is manually selected (14) and said first control device (48) transmits a corresponding selection signal (53) to the video system (42);

f) said video system (42) digitally stores second voice sound data portions (57) for said first talking character (18) corresponding to said first set of alternative verbal expressions (52);

g) said video system (42) generates (54) a signal (41) from one portion of said second voice sound data (57) in accordance with said transmitted selection signal (53) for output as second voice sounds of said first talking character (18);

h) said video system (42) digitally stores third voice sound data portions (58) for said second talking character (17) corresponding to said second voice sound data portions (57);

-19-

5 i) said video system (42) generates (56) a signal (41) for output as voice sounds of said second talking character (17) from the third voice sound data portion (74) that corresponds to the second voice sound data portion (73) from which the signal for said second voice sound was generated; and

10 j) said video system (42) transmits a second set of alternative guidance words or other preprogrammed verbal expressions (52) in digital form in accordance with said transmitted selection signal to a second manually operated control device (47) which displays
15 (13) the second set of alternative verbal expressions in order to be selected and generate a signal (53) for corresponding voice sounds.

-20-

2. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of a talking character (18) for display on a video screen (11) and generates a signal (41) representing the voice of the talking character,
- wherein digital data for background scenes and digital data for the animated character (18) and digitized voice data are each stored separately on a data storing medium (43) and read by said video system (42), said method comprising the following steps:
- a) said video system (42) digitally stores first voice sound data (57) for said talking character (18) and generates (54) a signal (41) from this first voice sound data for output as first voice sounds;
- b) said video system (42) transmits (72) a first set of alternative guidance words or other preprogrammed verbal expressions (52) in digital form to a first manually operated control device (48) which displays (13) the first set of alternative verbal expressions (52);
- c) one of said first set of verbal expressions is manually selected (14) and said first control device (48) transmits a corresponding first selection signal (53) to the video system (42);

-21-

d) said video system (42) digitally stores second voice sound data portions (57) for said talking character (18) corresponding to said first set of alternative verbal expressions (52);

5 e) said video system (42) generates (54) a signal (41) from one portion of said second voice sound data (57) in accordance with said first transmitted selection signal (53) for output as second voice sounds of said talking character (18);

10 f) said video system (42) transmits a second set of alternative guidance words or other preprogrammed verbal expressions (52) in digital form to a second manually operated control device (47) which displays the second set of alternative verbal expressions;

15 g) one of said second set of verbal expressions is manually selected (14) and said second control device (47) transmits a corresponding second selection signal (53) to the video system (42);

20 h) said video system (42) digitally stores third voice sound data portions (58) for said talking character (18) corresponding to said second set of alternative verbal expressions (52); and

i) said video system (42) generates (56) a signal (41) from one portion (74) of said third voice sound data in accordance with said second transmitted selection signal for output as third voice sounds of said talking character (18).

-22-

3. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of two or more talking characters (17,18) for display on
5 a video screen (11) and generates a signal (41) representing the respective voices of the talking characters,
- wherein digital data for background scenes and
10 digital data for each animated character (17,18) and digitized voice data are each stored separately on a data storing medium (43) and read by said video system (42), said method comprising the following steps:
- 15 a) digitally storing animated picture data for first (18) and second (17) talking characters;
- b) said video system (42) generates (51,55) a signal (44) from said animated picture data (70,71)
20 for display as animated pictures on said video screen (11);
- c) said video system (42) digitally stores first voice sound data (57) for said first talking character
25 (18) and generates (54) a signal (41) from this first voice sound data (57) for output as first voice sounds;

-23-

- d) said video system (42) transmits (72) a first set of alternative guidance words or other preprogrammed verbal expressions (52) in digital form to a first display means (13);
- 5 e) one of said first set of verbal expressions is manually selected (14) and a first control device (48) transmits a corresponding selection signal (53) to the video system (42);
- 10 f) said video system (42) digitally stores second voice sound data portions (57) for said first talking character (18) corresponding to said first set of alternative verbal expressions (52);
- 15 g) said video system (42) generates (54) a signal (41) from one portion of said second voice sound data (57) in accordance with said transmitted selection signal (53) for output as second voice sounds of said first talking character (18);
- 20 h) said video system (42) digitally stores third voice sound data portions (58) for said second talking character (17) corresponding to said second voice sound data portions (57);

-24-

- 5 i) said video system (42) generates (56) a signal (41) for output as voice sounds of said second talking character (17) from the third voice sound data portion (74) that corresponds to the second voice sound data portion (73) from which the signal for said second voice sound was generated; and
- 10 j) said video system (42) transmits a second set (59) of alternative guidance words or other preprogrammed verbal expressions (52) in digital form in accordance with said transmitted selection signal to a display means (13) in order to be selected and generate a signal for corresponding voice sounds.
- 15 4. The method of claims 1 to 2, wherein said first and second control devices are the same control device.
- 20 5. The method of claims 1 to 3, wherein the animated pictures of said talking animated characters are generated from digitized images.
- 25 6. The method of claims 1 to 3, wherein at least some of said voice sounds are altered to indicate that the altered voice sounds represent unspoken thoughts of one of said characters.

-25-

7. The method of claim 1, wherein step g) comprising generating a signal for said second voice sounds is delayed and performed between said steps i) and j).
8. The method of claims 1 and 3, wherein said second voice sounds say the words of said first selected verbal expression.
9. The method of claims 1 and 3, wherein said second voice sounds say words in a voice different than the voices of said first and second talking characters.

-26-

10. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of two or more talking characters (17, 18) for display on a video screen (11) and generates a signal (41) representing the respective voices of the talking characters,

wherein digital data for background scenes and digital data for each animated character (17,18) and digitized voice data are each stored separately on a data storing medium (43) and read by said video system (42), said method comprising the following steps:

- a) digitally storing first animated picture data (70) for a first (18) talking character;
- b) said video system (42) generates (51) a signal (44) from said animated picture data (70) for display as animated pictures of a first scene on said video screen (11);
- c) said video system (42) transmits (72) a first set of alternative guidance words or other preprogrammed verbal expressions (52) in digital form to a display means (13) during said first scene;

-27-

- d) one of said first set of verbal expressions is manually selected (14) and a control device (48) transmits a corresponding first selection signal (53) to the video system (42);
- 5 e) said video system (42) digitally stores first voice sound data portions (57) for said first talking character (18) corresponding to said first set of alternative verbal expressions (52);
- 10 f) said video system (42) generates (54) a signal (41) from one portion of said first voice sound data (57) in accordance with said transmitted first selection signal (53) for output as first voice sounds of said first talking character (18);
- 15 g) said video system (42) transmits (72) to a display means (13) a second set of alternative guidance words or other preprogrammed verbal expressions (52) in digital form corresponding to a set of alternative actions;
- 20 h) one of said second set of verbal expressions is manually selected (14) and a control device (48) transmits a corresponding second selection signal (53) to the video system (42) thereby selecting an action from said set of alternative actions;

-28-

- i) digitally storing second animated picture data (71) for a second talking character (17, 18, 30, 36) animated in accordance with said selected action;
- 5 j) said video system (42) generates (55) a signal (44) from said second animated picture data (71) for display as animated pictures of said second character in a second scene on said video screen (11);
- 10 k) said video system (42) digitally stores second voice sound data portions (58) for said second talking character corresponding to said second set of alternative verbal expressions (52);
- 15 l) said video system (42) generates (56) a signal (41) from one portion (74) of said second voice sound data portions (58) in accordance with said transmitted second selection signal for output as voice sounds of said second talking character; and
- 20 m) said video system (42) transmits a third set (59) of alternative guidance words or other preprogrammed verbal expressions (52) in digital form in accordance with said transmitted selection signal to a display means (13) in order to be selected and generate a
- 25 signal for corresponding voice sounds.
11. The method of claim 10, wherein said set of alternative actions includes hugging and kissing and wherein said first or second scenes show said two characters
- 30 hugging and kissing.

-29-

AMENDED CLAIMS

[received by the International Bureau
on 15 December 1993 (15.12.93); original claims 1-11 amended; new claims
12-16 added; (12 pages)]

1. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of two or more talking characters (17,18) for display on a video screen (11) and generates a signal (41) representing the respective voices of the talking characters, the words of one character being selected by a human player (12), said method comprising the following steps:
 - (a) digitally storing distinctive voices for at least two talking characters, each in the form of voice sound data (57,58) representing a plurality of sentences, phrases, word segments or phonemes;
 - (b) digitally storing a preprogrammed branching dialog (60,69) between a first animated character (17) and a second animated character (18), each branch comprising a plurality of alternative verbal expressions (52);
 - (c) generating a video signal (44) representing an image of said said first and second characters (17,18) for display on a video screen (11);
 - (d) displaying on a hand-held controller apparatus (47) a first set (13) of alternatively selectable verbal expressions, each corresponding to a branch in said dialog;

-30-

- (e) receiving from said hand-held controller (47) a manually initiated signal (53) representing a selected verbal expression in said first set (13) of verbal expressions, thereby selecting a branch in the dialog;
- 5 (f) digitally reading first voice sound data (73) that corresponds to said selected verbal expression (15) for the voice of said first character;
- 10 (g) generating (54) an audio signal (41) from said first voice sound data (73) representing the voice of said first character (17) speaking said selected verbal expression (15);
- 15 (h) digitally reading second voice sound data (74) for the voice of said second character (18) speaking the verbal expression (16) that follows next in the selected branch of the dialog;
- 20 (i) generating (56) an audio signal (41) from said second voice sound data (74) representing the voice of said second character (18); and
- 25 (j) displaying on said hand-held controller (47) a second set of alternatively selectable verbal expressions that follows next for said first character (17) in the selected branch of the dialog.
2. The method of claim 1, wherein one of said characters is shown hugging or kissing (Fig 3) the other character.
- 30

-31-

3. The method of claim 1, wherein said hand-held controller is a portable game system (46) with program cartridge.
4. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of two or more talking characters (17,18) for display on a video screen (11) and generates a signal (41) representing the respective voices of the talking characters, the words of one character being selected by a human player (12), said method comprising the following steps:
- (a) digitally storing distinctive voices for at least two talking characters, each in the form of voice sound data (57,58) representing a plurality of sentences, phrases, word segments or phonemes;
- (b) digitally storing a preprogrammed branching dialog (60,69) between a first animated character (17) and a second animated character (18), each branch comprising a plurality of alternative verbal expressions (52);
- (c) generating a video signal (44) representing an image of said said first and second characters (17,18) for display on a video screen (11);
- (d) displaying on a hand-held controller apparatus (47) a first set (13) of alternatively selectable verbal expressions, each corresponding to a branch in said dialog;

-32-

- (e) receiving from said hand-held controller (47) a manually initiated signal (53) representing a selected verbal expression in said first set (13) of verbal expressions, thereby selecting a branch in the dialog;
- 5 (f) digitally reading first voice sound data (73) that corresponds to said selected verbal expression (15) for the voice of said first character;
- 10 (g) generating (54) an audio signal (41) from said first voice sound data (73) representing the voice of said first character (17) speaking said selected verbal expression (15);
- 15 (h) changing said video signal (44) to indicate that the first character (17) is speaking to the second character (18);
- 20 (i) digitally reading second voice sound data (74) for the voice of said second character (18) speaking the verbal expression (16) that follows next in the selected branch of the dialog;
- 25 (j) generating (56) an audio signal (41) from said second voice sound data (74) representing the voice of said second character (18);
- (k) changing said video signal (44) to indicate that the second character (18) is speaking to the first character (17); and
- 30 (l) displaying on said hand-held controller (47) a second set of alternatively selectable verbal expressions that follows next for said first character (17) in the selected branch of the dialog.

-33-

5. The method of claim 4, wherein said video signal (44) is changed to show moving lips on the face of the first character (17) when said audio signal (41) is generated (54) from said first voice sound data (73); and
5 wherein said video signal (44) is changed to show moving lips on the face of the second character (18) when said audio signal (41) is generated (56) from said second voice sound data (74).
- 10 6. The method of claim 5, wherein said moving lips are lip-synchronized with the corresponding voices in said audio signal.
- 15 7. The method of claim 4, further comprising the step of altering either or both of said audio signals (41) so that the signal indicates unspoken thoughts of an on-screen character.
- 20 8. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of two or more on-screen talking characters (17,18) for display on a video screen (11) and generates a signal (41) representing the respective voices of the
25 on-screen talking characters and an off-screen character, the words of the off-screen character being selected by a human player (12), said method comprising the following steps:
- 30 (a) digitally storing distinctive voices for at least three talking characters, each in the form of voice sound data (57,58) representing a plurality of sentences, phrases, word segments or phonemes;

-34-

- 5 (b) digitally storing a preprogrammed branching dialog (60,69) between a first on-screen animated character (17) and a second on-screen character (18) and an off-screen character, each branch comprising a plurality of alternative verbal expressions (52);
- 10 (c) generating a video signal (44) representing an image of said first and second on-screen characters (17,18) for display on a video screen (11);
- 15 (d) displaying on a hand-held controller apparatus (48) a first set (13) of alternatively selectable verbal expressions, each corresponding to a branch in said dialog;
- 20 (e) receiving from said hand-held controller (48) a manually initiated signal (53) representing a selected verbal expression in said first set (13) of verbal expressions, thereby selecting a branch in the dialog;
- 25 (f) digitally reading first voice sound data that corresponds to said selected verbal expression for the voice of said off-screen character;
- 30 (g) generating (54) an audio signal (41) from said first voice sound data representing the voice of said off-screen character speaking said selected verbal expression;
- (h) digitally reading second voice sound data (73) for the voice of said first on-screen character (18) speaking the verbal expression (15) that follows next in the selected branch of the dialog;

-35-

(i) generating (54) an audio signal (41) from said second voice sound data (73) representing the voice of said first on-screen character (18);

5 (j) digitally reading third voice sound data (74) for the voice of said second on-screen character (17) speaking the verbal expression (16) that follows next in the selected branch of the dialog;

10 (k) generating (56) an audio signal (41) from said third voice sound data (74) representing the voice of said second on-screen character (17); and

15 (l) displaying on said hand-held controller (47) a second set of alternatively selectable verbal expressions that follows next for said off-screen character in the selected branch of the dialog.

9. The method of claim 8, wherein said video
20 signal (44) is changed to indicate which of said on-screen characters is speaking.

10. The method of claim 9, wherein said video
25 signal (44) is changed to show moving lips on the face of said first on-screen character (18) when said audio signal is generated from said second voice sound data (73) and to show moving lips on the face of said second on-screen character (17) when said audio signal is generated from said third voice sound data (74).

30 11. The method of claim 10, wherein said moving lips are lip-synchronized with the corresponding voices in said audio signal (41).

-36-

12. The method of claim 8, wherein said hand-held controller is a portable game system (46) with program cartridge.
- 5 13. The method of claim 8, wherein said selected verbal expression for said off-screen character is displayed on said video screen.
- 10 14. The method of claim 8, wherein one of said characters is shown hugging or kissing (Fig. 3) the other character.
- 15 15. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of two or more talking characters (17,18) for display on a video screen (11) and generates a signal (41) representing the respective voices of the talking characters, the words of the characters being
20 selected by two corresponding human players (10,12), said method comprising the following steps:
- 25 (a) digitally storing distinctive voices for at least two talking characters, each in the form of voice sound data (57,58) representing a plurality of sentences, phrases, word segments or phonemes;
- 30 (b) digitally storing a preprogrammed branching dialog (60,69) between a first animated character (18) and a second animated character (17), each branch comprising a plurality of alternative verbal expressions (52);

-37-

(c) generating a video signal (44) representing an image of said said first and second characters (18,17) for display on a video screen (11);

5 (d) displaying on a first hand-held controller apparatus (48) a first set (13) of alternatively selectable verbal expressions, each corresponding to a branch in said dialog;

10 (e) receiving from said first hand-held controller (48) a manually initiated signal (53) representing a first selected verbal expression in said first set (13) of verbal expressions, thereby selecting a branch in the dialog;

15 (f) digitally reading first voice sound data (73) that corresponds to said first selected verbal expression for the voice of said first character (18);

20 (g) generating (54) an audio signal (41) from said first voice sound data (73) representing the voice of said first character (18) speaking said first selected verbal expression;

25 (h) displaying on a second hand-held controller (47) apparatus a second set of alternatively selectable verbal expressions that follows next for said second character (17) in the selected branch of the dialog;

30 (i) receiving from said second hand-held controller (47) a manually initiated signal (53) representing a second selected verbal expression in said second set (26) of verbal expressions, thereby selecting a next branch in the dialog;

-38-

(j) digitally reading second voice sound data (74) that corresponds to said second selected verbal expression for the voice of said second character (17); and

5 (k) generating (56) an audio signal (41) from said second voice sound data (74) representing the voice of said second character (17) speaking said second selected verbal expression.

10 16. A method of simulating voice conversations by a video system (42) that generates a signal (44) representing animated pictures of two or more talking characters (17,18) for display on a video screen (11) and generates a signal (41)
15 representing the respective voices of the talking on-screen characters and two off-screen characters, the words of the off-screen characters being selected by two corresponding human players (10,12), said method comprising the following
20 steps:

(a) digitally storing distinctive voices for at least four talking characters, each in the form of voice sound data (57,58) representing a plurality of
25 sentences, phrases, word segments or phonemes;

(b) digitally storing a preprogrammed branching dialog (60,69) between a first on-screen character (18) and a second on-screen character (17) and a first
30 off-screen character and a second off-screen character, each branch comprising a plurality of alternative verbal expressions (52);

-39-

(c) generating a video signal (44) representing an image of said first and second on-screen characters (18,17) for display on a video screen (11);

5 (d) displaying on a first hand-held controller apparatus (48) a first set (13) of alternatively selectable verbal expressions, each corresponding to a branch in said dialog;

10 (e) receiving from said first hand-held controller (48) a manually initiated signal (53) representing a first selected verbal expression in said first set (13) of verbal expressions, thereby selecting a branch in the dialog;

15 (f) digitally reading first voice sound data (73) that corresponds to said first selected verbal expression for the voice of said first off-screen character;

20 (g) generating (54) an audio signal (41) from said first voice sound data (73) representing the voice of said first off-screen character speaking said first selected verbal expression;

25 (h) digitally reading second voice sound data (73) for the voice of said first on-screen character (18) speaking the verbal expression that follows next in the selected branch of the dialog;

30 (i) generating (56) an audio signal (41) from said second voice sound data (74) representing the voice of said first on-screen character;

-40-

5 (j) displaying on a second hand-held controller (47) apparatus a second set of alternatively selectable verbal expressions that follows next for said second off-screen character (18) in the selected branch of the dialog;

10 (k) receiving from said second hand-held controller (47) a manually initiated signal (53) representing a second selected verbal expression in said second set (26) of verbal expressions, thereby selecting a next branch in the dialog;

15 (l) digitally reading third voice sound data (73) that corresponds to said second selected verbal expression for the voice of said second off-screen character;

20 (m) generating (54) an audio signal (41) from said third voice sound data (73) representing the voice of said second off-screen character speaking said second selected verbal expression;

25 (n) digitally reading fourth voice sound data (74) for the voice of said second on-screen character (17) speaking the verbal expression that follows next in the selected branch of the dialog; and

30 (o) generating (56) an audio signal (41) from said fourth voice sound data (74) representing the voice of said second on-screen character (17).

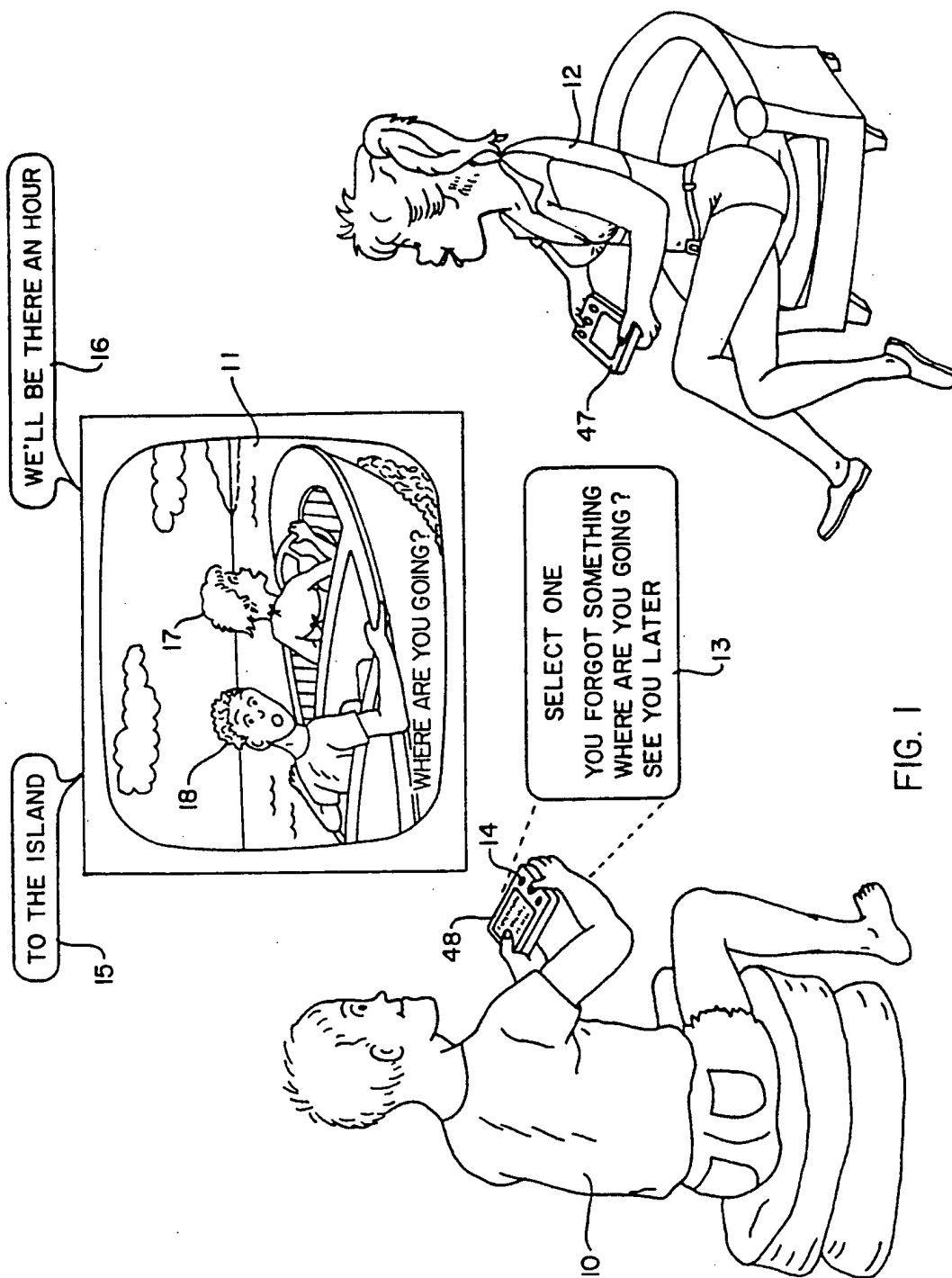


FIG. 1

2/10

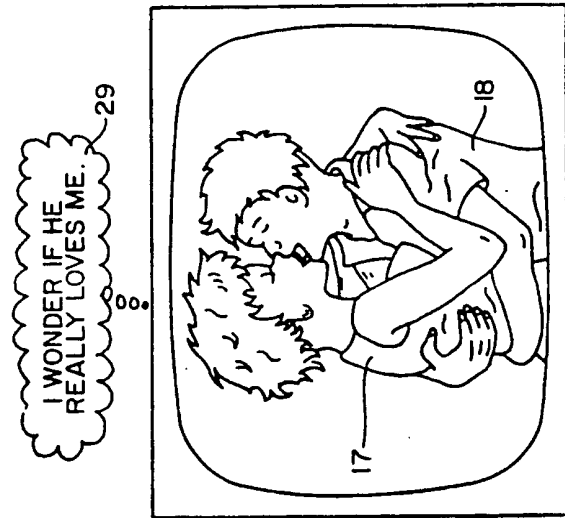


FIG. 3

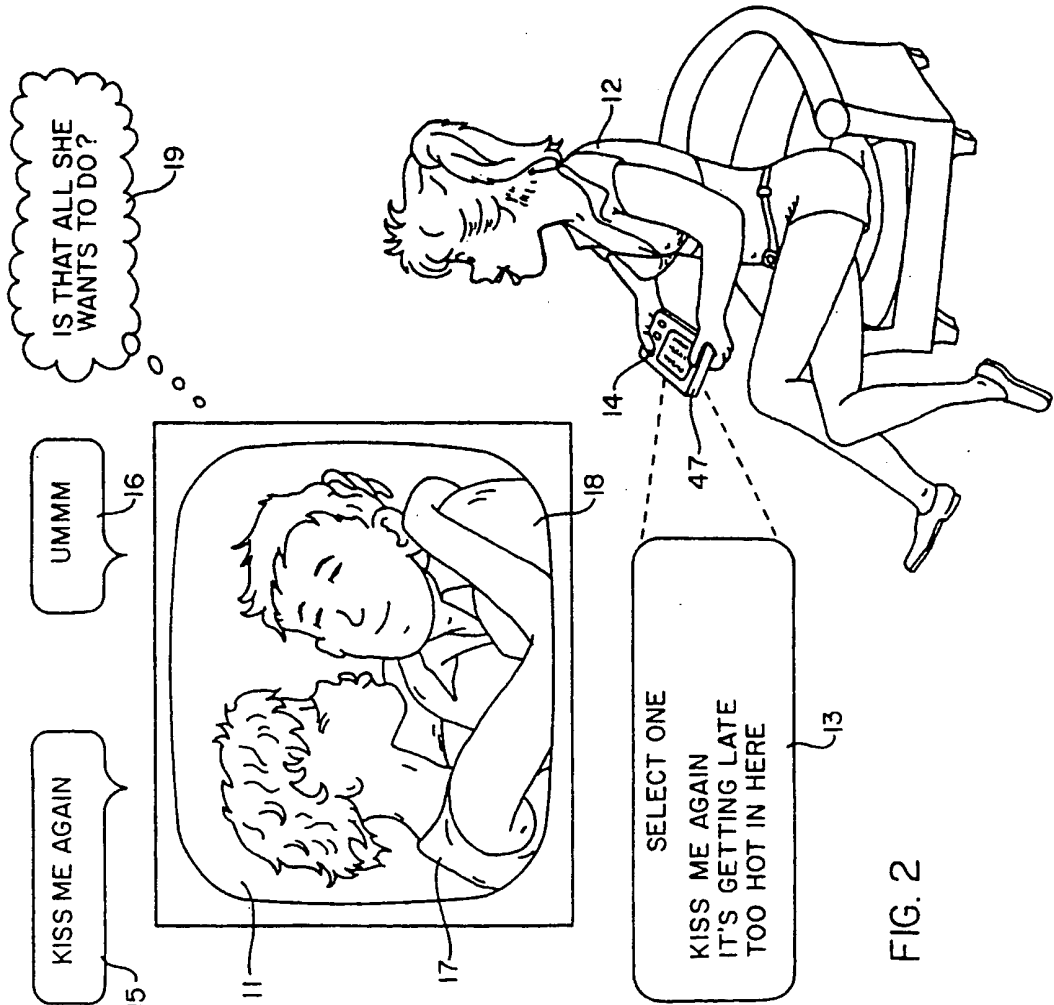


FIG. 2

SUBSTITUTE SHEET

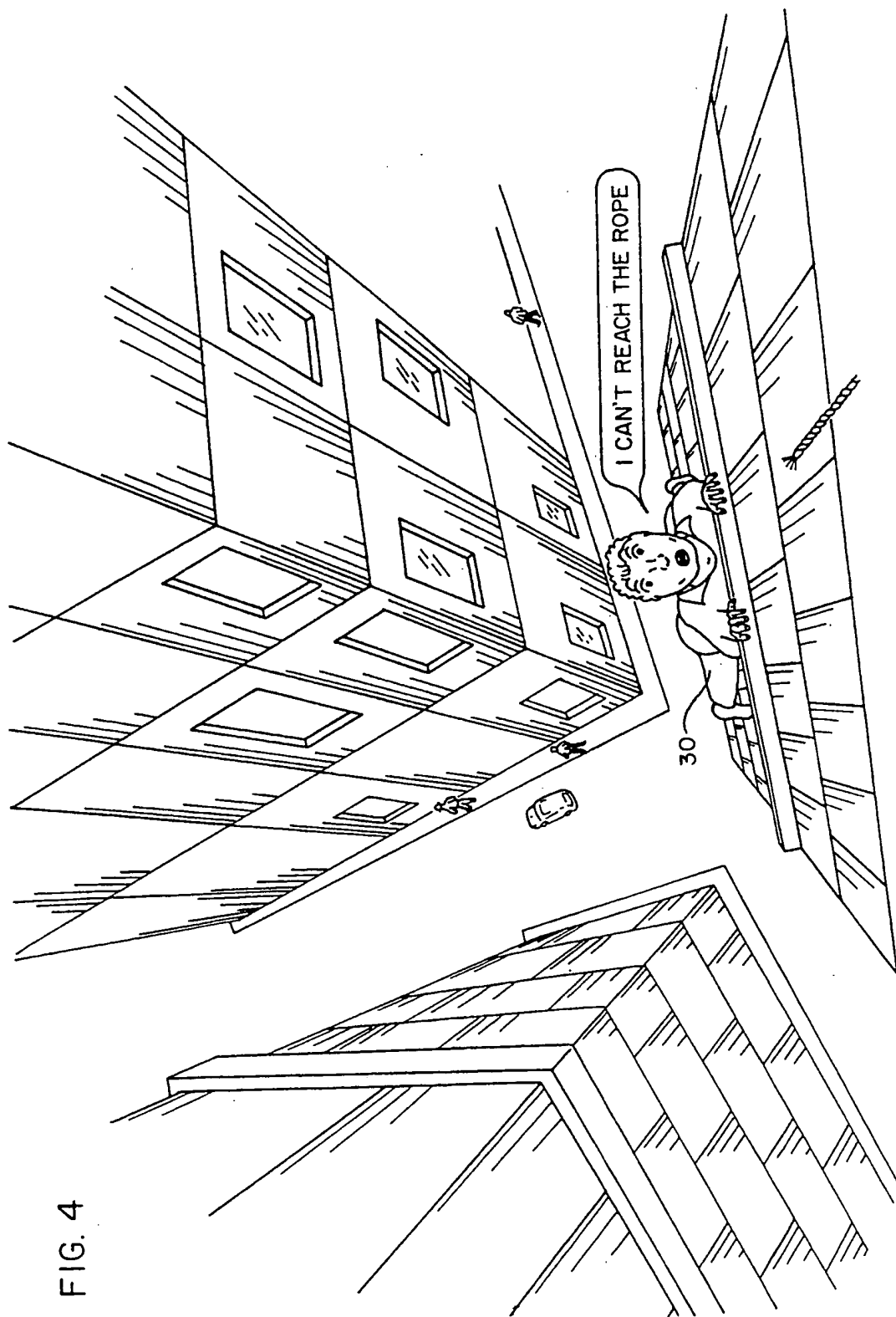


FIG. 4

4/10

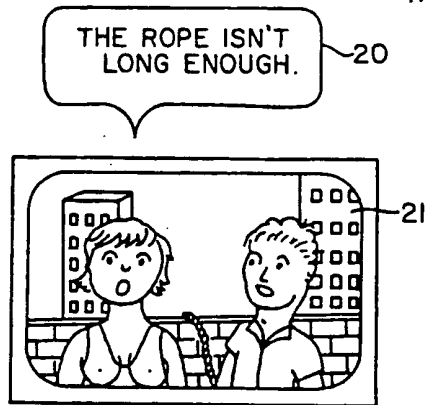
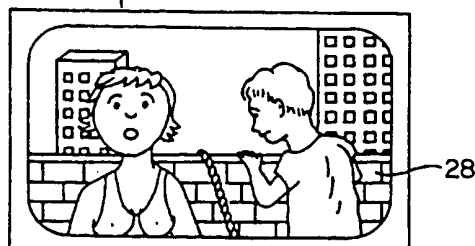
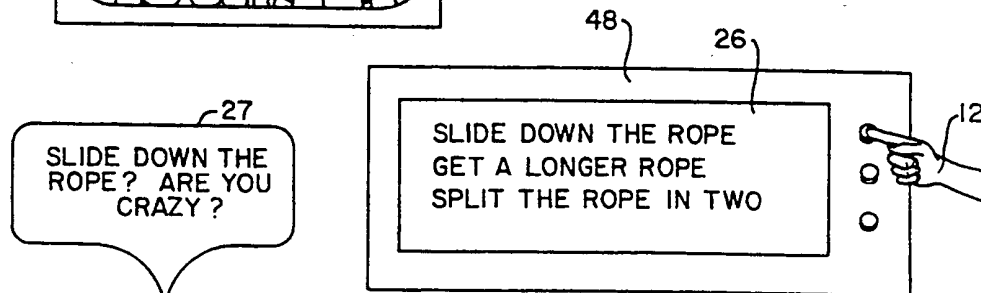
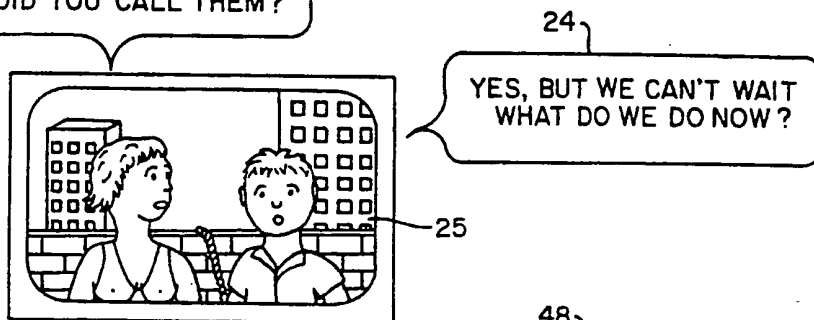
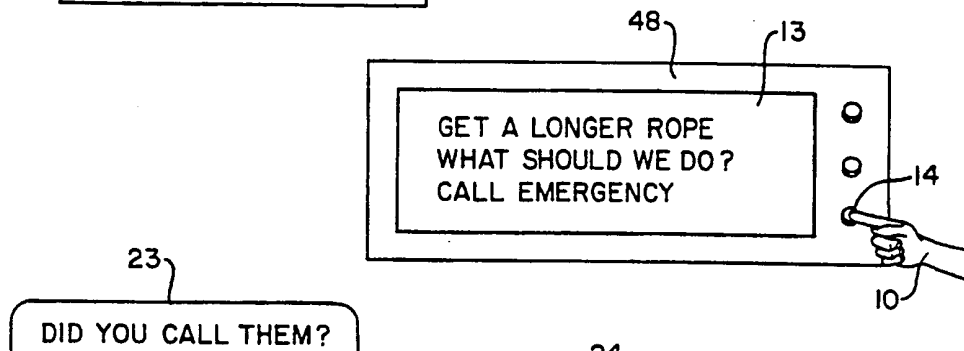
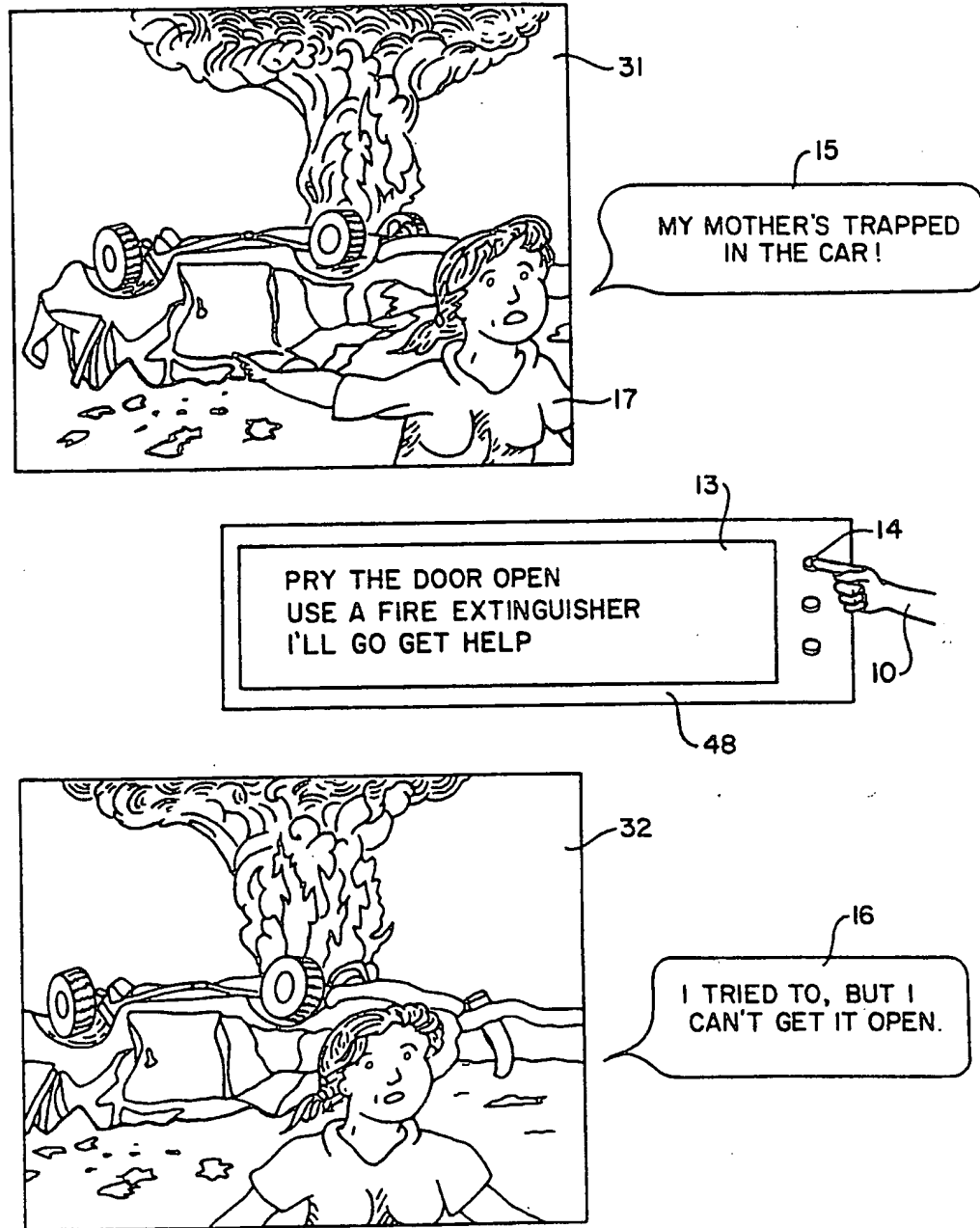


FIG. 5



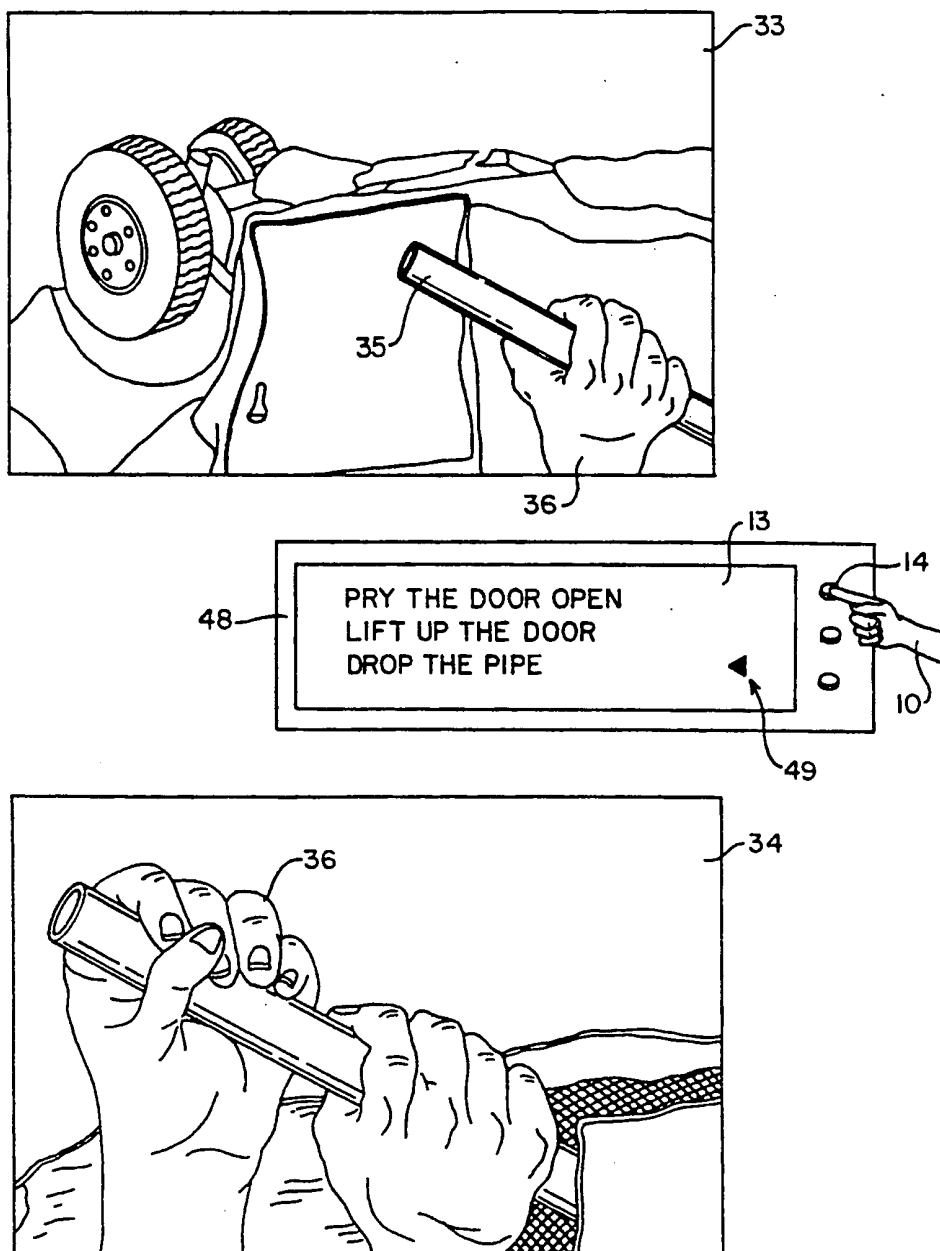
5/10

FIG. 6



SUBSTITUTE SHEET

FIG. 7



SUBSTITUTE SHEET

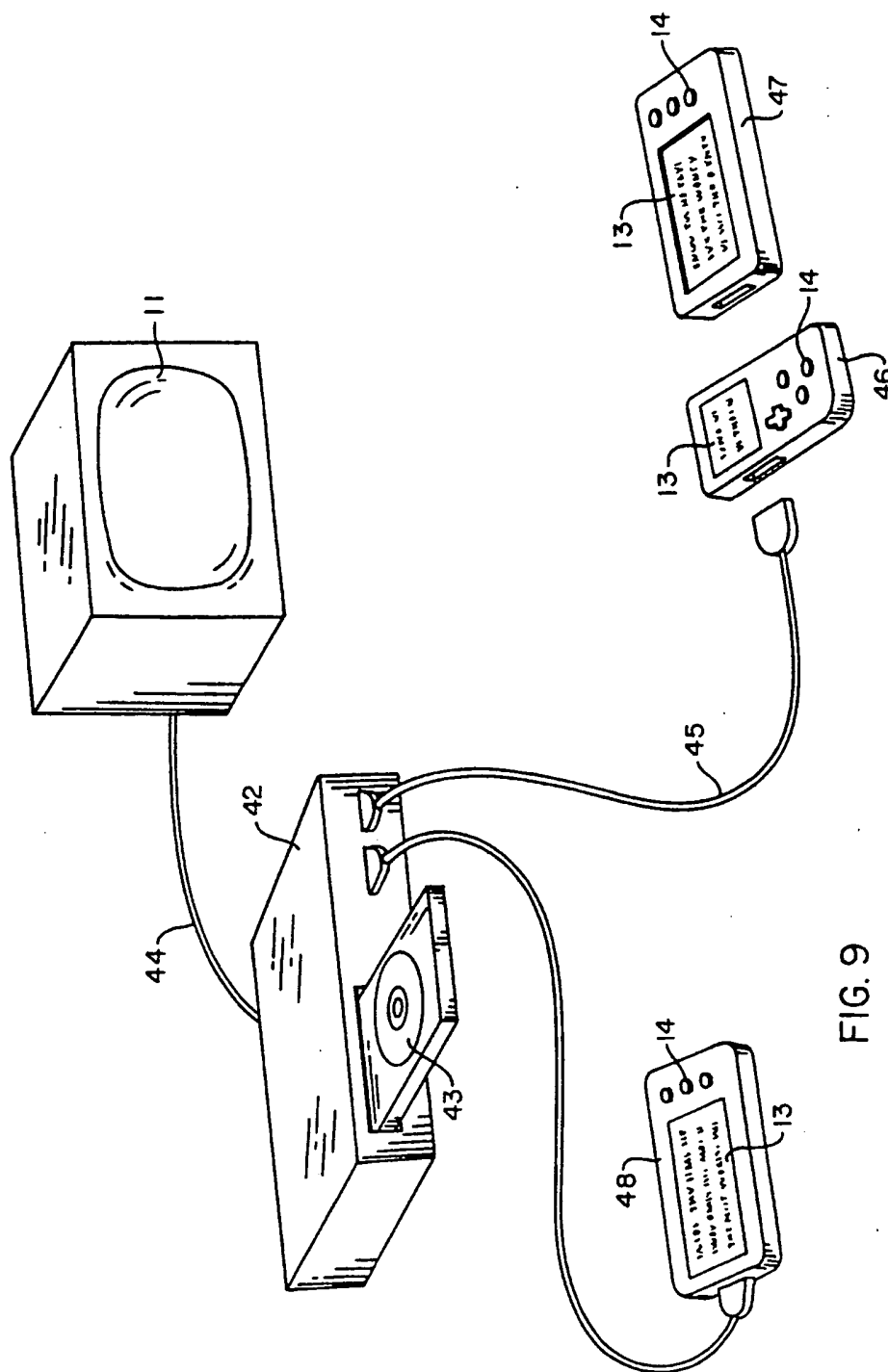


FIG. 9

9/10

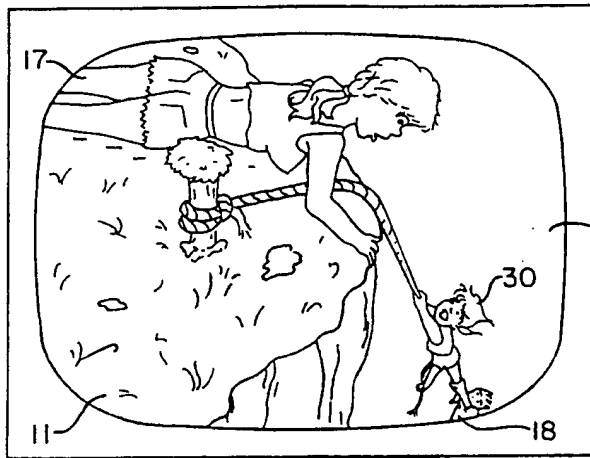
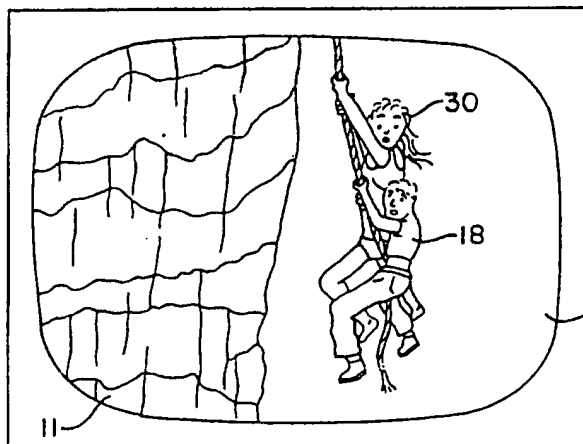
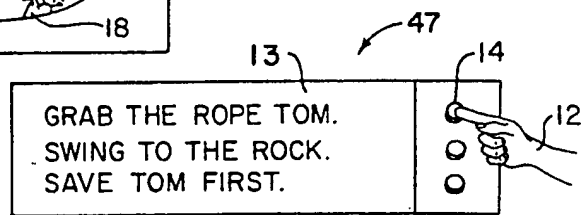
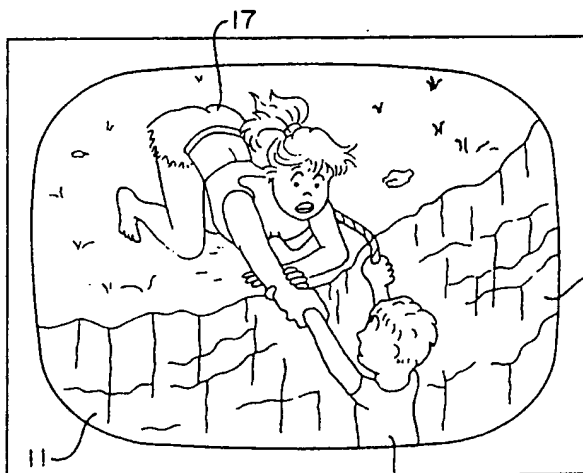


FIG. 10



GRAB THE ROPE TOM

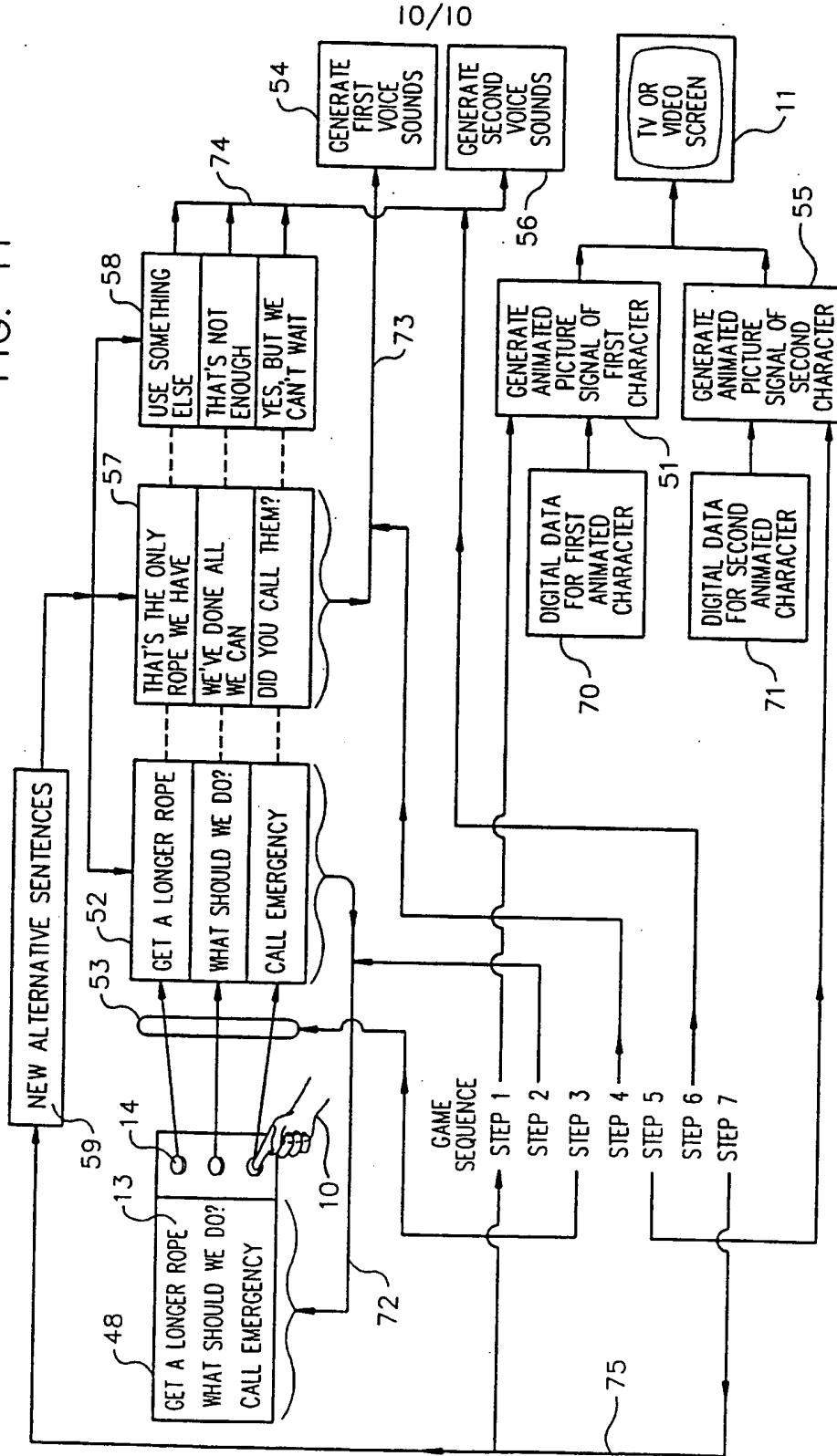
I'VE GOT IT



HOLD MY ARM

SUBSTITUTE SHEET

FIG. 11



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/05237**A. CLASSIFICATION OF SUBJECT MATTER**- IPC(5) : A63B 09/24
US CL : 273/434

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 273/434-437, Dig 28, 85G; 364/410; 434/323, 321

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
NoneElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)
None**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 4,445,187 (Best) 24 April 1984, see entire document.	1-10
Y	US, A, 5,006,987 (Harless) 09 April 1991, see particularly Fig 2, also see col. 5 through col. 6 line 68; and col. 8, lines 4-30.	1-11

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be part of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Z* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

02 September 1993

Date of mailing of the international search report

20 OCT 1993

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. NOT APPLICABLE

Authorized officer

FOR J. Harrison

Telephone No. (703) 308-2217

THIS PAGE BLANK (USPTO)